

BEST AVAILABLE COPY

Appendix A

Comparison of the amino acid sequences of the Cysteine proteases from soybean clone sr3c.pk003.d10:fs (SEQ ID NO:24) and *Phaseolus vulgaris* set forth NCBI General Identifier No. 2511691. Amino acids conserved among the two sequences are indicated with a single asterisk above the conserved residues. Dashes are used by the program to maximize alignment of the sequences. Residues conserved in all known cysteine proteases are underlined. Amino acids in the active site and cysteines in disulfide bridges are indicated by a dot and by a triangle below the conserved residues, respectively (Karrer et al.(1993) P.N.A.S.90, 3036-3067).

```

** * **** * * * * * * * * * * * * * * * * * *
SEQ ID NO: 24  MANLSLLFFGLLLFSA-AVATVERIDDEDNLLTRQVVPDAE-DRHLINAEHHFSAPKTKF
Gi:2511691     MARYTL--CALLLFAVAARAGGASTDADDILIRQVPEGEVEDHLINAEHHFSTPKSRF

*****
SEQ ID NO: 24  AKTYATQEEHHRFRIFFQNLLRAKSHQKLDPSAVHGVTRFSDLTPSEFRGQFLGLKPLR
Gi:2511691     GKTYATKEEHDHFRFGVFKSNMRRARLHAQLDPSAVHGVTRFSDLTPAEFRKFLGLKPLR

** ***** ** ***** ** * * * * * * * * * * * *
SEQ ID NO: 24  LPSDAQKAPILPTSDLPDWDHGAIVGVKNGSGCWCWSPSAVGALEGARFLSTGGL
Gi:2511691     LPAHAQRAPILPTMNLFPKDFWRDKGAVINVKDQSGSCGSCWSTTGGALEGARFLATGEL
                *   ▲   •

*****
SEQ ID NO: 24  VSLSEQQLVDCDEECDFEERGACDSGCNGGLMTTAFETLKAGGLMREEDYPYTGDRGP
Gi:2511691     VSLSEQQLVDCDHVCDPEEYGS CDSGCNGGLMNAFEYTLIGSGGVQREKDYPTTGRD-GT
                ▲

*****
SEQ ID NO: 24  CKFDKSKLAASVANPFSVSLDEEQIAANLVKNGPLAVGINAVFMQTYIGGVSCFYICGRH
Gi:2511691     CKFDKSKIAASVSNYSVLSLDEEQIAANLVKNGPLAVAINAVFMQTYVGVSCFYICGRH
                ▲

*****
SEQ ID NO: 24  LDHGVLVVGYSAGAYAPIRFKKPYWIIKNSWGESWGEZGYIKICRGRNVCGVDSMVSTV
Gi:2511691     LDHGVLVVGYSAGAYAPIRFKKPYWIIKNSWGENWGGNGYIKICRGRNVCGVDSMVSTV
                *   •   ▲

***
SEQ ID NO: 24  AAIHVSNH
Gi:2511691     GATHASTQ
```